GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: June 24, 1999, 01:22:16 ; Search time 67.43 Seconds
(without alignments)
410.133 Million cell updates/sec

Title: US-09-205-015-2
Perfect score: 147
Sequence: 1 agataactgggccaaccatg.....cctccccctcctgtttatct 147

Scoring table: IDENTITY_NUC

Searched: 240622 seqs, 94065609 residues

Database : N_Geneseq_34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 27 | 18.4 | 1420 | 1 | V32475 | Bovine retinaldehy |
| 2 | 26.8 | 18.2 | 2158 | 1 | Q25156 | Alpha-GalNAc from |
| 3 | 26.8 | 18.2 | 1840 | 1 | Q81826 | Alpha-N-acetylgala |
| c 4 | 26.8 | 18.2 | 334 | 1 | T26078 | Human gene signatu |
| 5 | 26.4 | 18.0 | 495 | 1 | N91109 | Human reg cDNA., N |
| 6 | 26.4 | 18.0 | 441 | 1 | Q05622 | Sequence encoding |
| 7 | 26.4 | 18.0 | 498 | 1 | N81962 | Sequence of human |
| c 8 | 26.4 | 18.0 | 6327 | 1 | T32301 | Dermatomyositis sp |
| 9 | 26.2 | 17.8 | 1827 | 1 | T09865 | First intron promo |
| c 10 | 26 | 17.7 | 142 | 1 | V41229 | Prevotella nigresc |
| c 11 | 25.8 | 17.6 | 2525 | 1 | V02016 | DNA encoding tumou |
| 12 | 25.6 | 17.4 | 383 | 1 | Q61216 | Human brain Expres |
| 13 | 25.6 | 17.4 | 1254 | 1 | Q58718 | Human gamma-1 chai |
| c 14 | 25.4 | 17.3 | 1371 | 1 | Q47076 | B. canis 21B4/rhop |
| 15 | 25.2 | 17.1 | 565 | 1 | N70930 | Sequence encoding |
| 16 | 25.2 | 17.1 | 495 | 1 | Q40477 | MUREG-1 coding reg |
| c 17 | 25.2 | 17.1 | 3969 | 1 | Q49757 | pTK gene HpTK5. Ne |
| c 18 | 25.2 | 17.1 | 3969 | 1 | T03099 | Protein tyrosine-k |
| c 19 | 25.2 | 17.1 | 4290 | 1 | Q92641 | Human non-differen |
| c 20 | 25.2 | 17.1 | 4290 | 1 | T18394 | Receptor type tyro |
| c 21 | 25.2 | 17.1 | 4290 | 1 | T42593 | Coding sequence fo |
| c 22 | 25.2 | 17.1 | 4290 | 1 | T51235 | Receptor-type tyro |
| c 23 | 25.2 | 17.1 | 1228 | 1 | T94471 | Human Fchd605 gene |
| c 24 | 25 | 17.0 | 3252 | 1 | Q54375 | Human BDNF vector. |
| c 25 | 25 | 17.0 | 3370 | 1 | Q79534 | Bovine tracheal an |
| 26 | 25 | 17.0 | 6511 | 1 | Q95493 | Human Cdn-2 DNA. N |
| 27 | 25 | 17.0 | 2072 | 1 | Q95492 | Human Cdn-1 cDNA. |
| 28 | 25 | 17.0 | 1968 | 1 | T17375 | Bcl-Y cDNA. New is |
| 29 | 25 | 17.0 | 2094 | 1 | T42138 | Bak gene. Screenin |
| c 30 | 25 | 17.0 | 2460 | 1 | V44303 | Human secreted pro |
| 31 | 25 | 17.0 | 6623 | 1 | T75251 | Nucleotide sequenc |
| 32 | 25 | 17.0 | 2094 | 1 | V61498 | Bak cDNA. New Bak- |
| c 33 | 24.8 | 16.9 | 1843 | 1 | T38308 | Human B-cell trans |
| c 34 | 24.6 | 16.7 | 2068 | 1 | Q13224 | DNA encoding Termi |
| c 35 | 24.6 | 16.7 | 1758 | 1 | V58786 | Human phospholipas |
| c 36 | 24.4 | 16.6 | 1621 | 1 | Q66423 | B. campestris Bgp1 |
| c 37 | 24.4 | 16.6 | 3300 | 1 | V31822 | Mutant Aspergillus |
| 38 | 24.4 | 16.6 | 1855 | 1 | V68998 | DNA molecule encod |
| c 39 | 24.2 | 16.5 | 2035 | 1 | N70687 | DNA encoding human |
| c 40 | 24.2 | 16.5 | 2676 | 1 | Q14850 | Clone pTB1284 enco |
| 41 | 24.2 | 16.5 | 387 | 1 | Q35904 | Anti-CD4 V-lambda |
| c 42 | 24.2 | 16.5 | 1514 | 1 | Q68267 | Maize 2-acyltransf |
| c 43 | 24.2 | 16.5 | 2131 | 1 | Q88155 | Human lung tumour |
| 44 | 24.2 | 16.5 | 816 | 1 | Q98724 | Anti-HIV-1 MAb 447 |
| 45 | 24.2 | 16.5 | 2481 | 1 | V33510 | Human G-protein ad |

ALIGNMENTS

RESULT 1
V32475
ID V32475 standard; cDNA; 1420 BP.
AC V32475;
DT 11-SEP-1998 (first entry)
DE Bovine retinaldehyde binding protein cDNA.
KW Bovine retinaldehyde binding protein; retinal pigment epithelium; RPE;
KW 11-cis-retinal; all-trans retinal; visual system; binding assay;
KW chromophore; ss.
OS Bos sp.
FH Key Location/Qualifiers
FT CDS 17. .812
FT /*tag= a
FT /product= "Retinaldehyde binding protein"
PN US5763578-A.
PD 09-JUN-1998.
PF 16-DEC-1994; 358171.
PR 16-DEC-1994; US-358171.
PA (FONG/) FONG H K W.
PI Fong HKW;
DR WPI; 98-347415/30.
DR P-PSDB; W48857.
PT Human and bovine retin-aldehyde-binding proteins - used to detect
PT aberration(s) of retinal binding in visual excitation systems
PS Disclosure; Fig 1; 39pp; English.
CC The present sequence represents the bovine retinaldehyde binding
CC protein cDNA isolated from a bovine retinal pigment epithelium (RPE)
CC cDNA library. The bovine retinaldehyde binding protein binds both
CC 11-cis-retinal and all-trans retinal. The invention claims that
CC molecular aberration of the visual system can be detected in binding
CC assays by observing any changes in the binding of the retinaldehyde
CC binding protein to its chromophores. The retinaldehyde binding protein
CC can also be used to raise antibodies, which in turn can be used to
CC detect changes of the protein in samples.
SQ Sequence 1420 BP; 291 A; 447 C; 384 G; 298 T;

Query Match 18.4%; Score 27; DB 1; Length 1420;
Best Local Similarity 62.7%; Pred. No. 6.3;
Matches 42; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
Qy       7 ctgggccaaccatgactcagtgcttctggaggccaacaggacttctgagtcatcctgtgg 66
           ||||| | || ||||||  | ||  |||| | ||  | | |||   || |  ||| | ||
Db     463 CTGGGGCCACTATGACTATGAGCCGCTGGGGACCTGCTGCACTCTGGACTATTCCAGGGG 522

Qy      67 gggtgga 73
           ||  ||
Db     523 GGACAGA 529
```

RESULT 2
Q25156
ID Q25156 standard; cDNA; 2158 BP.
AC Q25156;
DT 18-NOV-1992 (first entry)
DE Alpha-GalNAc from pAGB-3.
KW Lysosome; Schindler disease; infantile neuroaxonal dystrophy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 345. .1580
FT /*tag= a
FT /label= alpha-GalNAc
FT signal_peptide 345. .395
FT /*tag= b
FT mat_peptide 396. .1580
FT /*tag= c

```
FT   polya_signal    2073. .2078
FT                   /*tag=  d
FT   protein_bind    2025. .2029
FT                   /*tag=  e
FT                   /note= "recognised by the U4 small nuclear
FT                   ribonucleoprotein"
PN   WO9207936-A.
PD   14-MAY-1992.
PF   23-OCT-1991; U07872.
PR   24-OCT-1990; US-602608.
PA   (MOUN ) MOUNT SINAI SCHOOL MEDICINE.
PI   Bishop DF, Desnick RJ, Ioannou YA, Wang AM;
DR   WPI; 92-183672/22.
DR   P-PSDB; R24291.
PT   Cloning and expression of alpha-n-acetyl-galactose aminidase -
PT   used in enzyme replacement therapy for Schindler disease
PS   Disclosure; Fig 2 (A-D); 71pp; English.
CC   The sequence is of the pAGB-3 cDNA insert contg. the complete coding
CC   region for human alpha-GalNAc.
CC   The availability of the full length cDNA for alpha-GalNAc allows
CC   the study of the genomic organisation and evolution of this
CC   lysosomal gene, and the characterisation of molecular lesions
CC   causing Schindler disease.
SQ   Sequence   2158 BP;    517 A;    610 C;    576 G;    455 T;

  Query Match             18.2%;  Score 26.8;  DB 1;  Length 2158;
  Best Local Similarity   55.3%;  Pred. No. 8.2;
  Matches   52;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy      32 ctggaggccaacaggacttctgagtcatcctgtgggggtggaggtgggacaagggaaagg 91
           ||| |  ||| || ||     |||  |||     | | |||| ||||||| | || |||
Db    1542 CTGTATCCCATCAAGAACCTGGAGATGTCCCAGCAGTGAGGAGCTGGGACATGTGACAGG 1601

Qy      92 ggtgaatggtactgctgattacaacctctggtgc 125
              ||| ||  ||||    |  |  ||| ||
Db    1602 CTGTGGTGGCACCACTGAGCCTAGACCATGGAGC 1635


RESULT   3
Q81826
ID   Q81826 standard; cDNA; 1840 BP.
AC   Q81826;
DT   10-MAR-1995  (first entry)
DE   Alpha-N-acetylgalactosaminidase.
KW   Alpha-N-acetylgalactosaminidase; expression; enzyme;
KW   erythrocyte; amplification; primer; polymerase chain reaction; PCR;
KW   probe; blood; type A; type B; type AB; type O; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             73. .1308
FT                   /*tag=  a
PN   WO9411518-A.
PD   26-MAY-1994.
PF   08-NOV-1993; U10794.
PR   18-NOV-1992; US-977945.
PA   (GEMV ) GENENCOR INT INC.
PI   Berka RM;
DR   WPI; 94-183517/22.
DR   P-PSDB; R69101.
PT   Prodn. of human placental alpha-N-acetylgalactosaminidase - by
PT   expression in transformed host cells, used to convert type A, B
PT   or AB erythrocytes to type O
PS   Claim 13; Fig 2; 66pp; English.
CC   The two primers given in Q81823-24 were used in the screening of
CC   libraries contg. sequences specific for alpha-N-
CC   acetylgalactosaminidase cDNA clones.  A 466 bp fragment was
CC   obtained, which was then subcloned and used as a probe.
CC   Another probe given in Q81825 was used in a secondary
CC   screening process.  A full length alpha-N-acetylgalactosaminidase
CC   cDNA was obtained (see Q81826).
SQ   Sequence   1840 BP;    417 A;    539 C;    485 G;    399 T;

  Query Match             18.2%;  Score 26.8;  DB 1;  Length 1840;
  Best Local Similarity   55.3%;  Pred. No. 7.9;
  Matches   52;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy      32 ctggaggccaacaggacttctgagtcatcctgtgggggtggaggtgggacaagggaaagg 91
           ||| |  ||| || ||     |||  |||     | | |||| ||||||| | || |||
Db    1270 CTGTATCCCATCAAGAACCTGGAGATGTCCCAGCAGTGAGGAGCTGGGACATGTGACAGG 1329

Qy      92 ggtgaatggtactgctgattacaacctctggtgc 125
              ||| ||  ||||    |  |  ||| ||
Db    1330 CTGTGGTGGCACCACTGAGCCTAGACCATGGAGC 1363


RESULT   4
T26078/c
ID   T26078 standard; cDNA to mRNA; 334 BP.
AC   T26078;
DT   16-OCT-1996  (first entry)
DE   Human gene signature HUMGS08314.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1996; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   334 BP;    65 A;    98 C;    79 G;    79 T;

  Query Match             18.2%;  Score 26.8;  DB 1;  Length 334;
  Best Local Similarity   63.5%;  Pred. No. 5;
  Matches   40;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy      37 ggccaacaggacttctgagtcatcctgtgggggtggaggtgggacaagggaaaggggtga 96
           || ||    ||||  ||   |||||| ||| ||      |||||||||| |||| |||
Db     127 GGACACTCTGACTATTGCACAATCCTGGGGGAGTNCCCCAGGGACAAGGGGAAGGAGTGT 68

Qy      97 atg 99
           ||
Db      67 GTG 65


RESULT   5
N91109
ID   N91109 standard; DNA; 495 BP.
AC   N91109;
DT   21-JUN-1990  (first entry)
```

```
DE   Human reg cDNA.,
KW   reg proteins; islet cells; diabetes; insulin; ds.
OS   Homo sapiens.
PN   EP-303233-A.
PD   15-FEB-1989.
PF   9-AUG-1988; 112942.
PR   10-AUG-1987; JP-200514.
PA   (SHIO) Shionogi KK.
PI   Okamoto H, Itoh T, Teraoka H, Tsuzuki H, Yoshida N;
DR   WPI; 89-048048/07.
DR   P-PSDB; P94614.
PT   New human reg proteins -
PT   useful for regenerating islet B cells in diabetes treatment.
PS   Claim 1; Fig 1; 19pp; English.
CC   Gene encodes reg protein useful in regeneration of human pancreatic islet
CC   B cells in the treatment of diabetes.
SQ   Sequence   495 BP;   115 A;   137 C;   128 G;   115 T;


  Query Match             18.0%;  Score 26.4;  DB 1;  Length 495;
  Best Local Similarity   57.1%;  Pred. No. 7.4;
  Matches   48;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      49 ttctgagtcatcctgtgggggtggaggtgggacaagggaaaggggtgaatggtactgctg 108
           | || |  ||  | | ||| |     ||||    |  ||    || || ||| |||| ||
Db     221 TGCTCACCCAGGCCGAGGGTGCCTTTGTGGCCTCACTGATTAAGGAGAGTGGCACTGATG 280

Qy     109 attacaacctctggtgctgcctcc 132
           | | |||  |||||    ||||||
Db     281 ACTTCAATGTCTGGATTGGCCTCC 304


RESULT   6
Q05622
ID   Q05622 standard; DNA; 441 BP.
AC   Q05622;
DT   07-JAN-1991  (first entry)
DE   Sequence encoding reg protein analogue.
KW   Reg protein; diabetes; B cells; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             1. .441
FT                   /*tag=  a
FT                   /label=reg protein analogue
PN   EP-383453-A.
PD   22-AUG-1990.
PF   30-JAN-1990; 300963.
PR   30-JAN-1989; JP-022132.
PA   (SHIO ) SHIONOGI SEIYAKU KK.
PI   Okamoto H, Itoh T, Teraoka H, Tsuzuki H, Yoshida N;
DR   WPI; 90-255705/34.
DR   P-PSDB; R06425.
PT   New human reg. protein and gene encoding it - used for
PT   stimulating and activating pancreatic B cells
PS   Disclosure; fig 1; 15pp; English.
CC   The reg protein encoded by this sequence comprises residues 20 (Gln)
CC   to 165 (Asn) of the human reg protein, opt. preceded by an N-Met
CC   residue.  It is involved in regeneration of insulin-producing
CC   pancreatic B cells and hence is used in the treatment of diabetes.
SQ   Sequence   441 BP;    106 A;    119 C;    118 G;    98 T;


  Query Match             18.0%;  Score 26.4;  DB 1;  Length 441;
  Best Local Similarity   57.1%;  Pred. No. 7.2;
  Matches   48;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      49 ttctgagtcatcctgtgggggtggaggtgggacaagggaaaggggtgaatggtactgctg 108
           | || |  ||  | | ||| |     ||||    |  ||    || || ||| |||| ||
Db     167 TGCTCACCCAGGCCGAGGGTGCCTTTGTGGCCTCACTGATTAAGGAGAGTGGCACTGATG 226

Qy     109 attacaacctctggtgctgcctcc 132
           | | |||  |||||    ||||||
Db     227 ACTTCAATGTCTGGATTGGCCTCC 250


RESULT   7
N81962
ID   N81962 standard; DNA; 498 BP.
AC   N81962;
DT   02-FEB-1991  (first entry)
DE   Sequence of human reg cDNA
KW   Pancreatic islet B cell regeneration; diabetes; therapy; ss.
OS   Homo sapiens.
FH   Key          Location/Qualifiers
FT   cds                  CDS   1. .498
FT                   /tag=a
PN   EP-286114-A.
PD   12-DEC-1988.
PF   08-APR-1988; 105623.
PR   10-AUG-1987; JP-200514.
PA   (SHIO) Shionogi Seiyaku Kk.
PI   Okamoto H;;
DR   WPI; 88-287314/41.
DR   P-PSDB; P81514.
PT   Rat and human reg genes -
PT   used for producing proteins for regeneration of
PT   insulin-producing B cells of patients with diabetes
PS   Claim 1; fig 3; 12pp; English.
CC   The reg gene is specifically expressed in regenerating pancreatic islet
CC   B cells.  A gene hybridising to a probe corresponding to at least a part
CC   of the whole base sequence of rat reg gene or human reg gene is claimed.
CC   By mass producing the proteins encoded by the gene it may be possible
CC   to open a new dimension in the treatment of diabetes
SQ   Sequence   498 BP;   116 A;   136 C;   130 G;   116 T;


  Query Match             18.0%;  Score 26.4;  DB 1;  Length 498;
  Best Local Similarity   57.1%;  Pred. No. 7.4;
  Matches   48;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      49 ttctgagtcatcctgtgggggtggaggtgggacaagggaaaggggtgaatggtactgctg 108
           | || |  ||  | | ||| |     ||||    |  ||    || || ||| |||| ||
Db     224 TGCTCACCCAGGCCGAGGGTGCCTTTGTGGCCTCACTGATTAAGGAGAGTGGCACTGATG 283

Qy     109 attacaacctctggtgctgcctcc 132
           | | |||  |||||    ||||||
Db     284 ACTTCAATGTCTGGATTGGCCTCC 307


RESULT   8
T32301/c
ID   T32301 standard; cDNA; 6327 BP.
AC   T32301;
DT   30-OCT-1996  (first entry)
DE   Dermatomyositis specific autoantigen, Mi-2, coding sequence.
KW   Mi-2; autoantigen; collagen disease; chromosome 12; 12p13;
KW   helicase; dermatomyositis; diagnosis; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             1. .5738
FT                   /*tag=  a
FT                   /product= Mi-2
FT                   /note= "the first ATG is at nucleotide 90, the sequence
FT                   given starts at nucleotide 91"
FT   3'utr           1579. .6417
FT                   /*tag=  b
FT   polya_signal    6234. .6240
FT                   /*tag=  c
PN   DE19509279-C1.
PD   15-MAY-1996.
PF   15-MAR-1995; 009279.
PR   15-MAR-1995; DE-009279.
PA   (PRIV-) PRIVATES INST IMMUNOLOGIE & MOLEKULARGEN.
PI   Renz M,  Seelig HP;
DR   WPI; 96-240280/25.
DR   P-PSDB; R99534.
```

```
PT   DNA encoding dermatomyositis specific auto:antigen - useful for
PT   differential diagnosis and treatment of dermatomyositis
PS   Claim 3; Fig 2; 20pp; German.
CC   The present sequence encodes a 218 kD dermatomyositis specific auto-
CC   antigen, designated Mi-2. The sequence numbering given in the
CC   specification starts at nucleotide 91, i.e. the 5' UTR is omitted, as
CC   well as the 'A' of the first ATG start codon. The protein is hydrophilic,
CC   acidic and protruding regions of the protein are characteristic of
CC   helicases. The gene corresponding to this cDNA was localised to
CC   chromosome 12 (12p13). The DNA can be used for the recombinant production
CC   of Mi-2 which is used for, e.g. the differential diagnosis of collagen
CC   diseases, esp. dermatomyositis, e.g. by immunoassay or Western blotting.
SQ   Sequence   6327 BP;    1761 A;    1473 C;    1740 G;    1353 T;


  Query Match             18.0%;  Score 26.4;  DB 1;  Length 6327;
  Best Local Similarity   75.0%;  Pred. No. 15;
  Matches   33;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy     66 ggggtggaggtgggacaagggaaaggggtgaatggtactgctga 109
          |||| | || |||||||||| ||||| |  ||| |  |||||| |
Db   5810 GGGGAGAAGCTGGGACAAGAGAAAGTGAGGAAGGTCACTGCTCA 5767


RESULT   9
T09865
ID   T09865 standard; DNA; 1827 BP.
AC   T09865;
DT   02-JAN-1997  (first entry)
DE   First intron promoter region of p75 tumour necrosis factor receptor.
KW   Tumour necrosis factor receptor; TNF; TNF-R; sequence motif; probe;
KW   motif region; inhibition; induction; inflammation; septic shock;
KW   rheumatism; autoimmune disease; graft vs. host reaction; promoter;
KW   ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   repeat_region   238. .246
FT                   /*tag= a
FT                   /label= TCC repeat.
FT   misc_feature    255. .264
FT                   /*tag= b
FT                   /label= NF-kB.
FT   misc_feature    288. .296
FT                   /*tag= c
FT                   /label= CK-1.
FT   misc_feature    388. .393
FT                   /*tag= d
FT                   /label= IFN.
FT   misc_feature    422. .430
FT                   /*tag= e
FT                   /label= CK-1.
FT   misc_feature    567. .572
FT                   /*tag= f
FT                   /label= IFN.
FT   tata_signal     630. .635
FT                   /*tag= g
FT   misc_feature    656. .672
FT                   /*tag= h
FT                   /label= YY1, CAP-site.
FT   misc_feature    709. .715
FT                   /*tag= i
FT                   /label= CK-2.
FT   misc_feature    757. .762
FT                   /*tag= j
FT                   /label= IFN.
FT   repeat_region   1310. .1324
FT                   /*tag= k
FT                   /label= TCC repeat.
FT   misc_feature    1780. .1795
FT                   /*tag= l
FT                   /label= 3' splice site.
FT   mat_peptide     1795
FT                   /*tag= m
FT                   /note= "Fragment of mature peptide
FT                   (start of exon 2).
PN   WO9531206-A1.
PD   23-NOV-1995.
PF   11-MAY-1995; U05853.
PR   11-MAY-1994; IL-109633.
PA   (WEIN/) WEINWURZEL H.
PA   (YEDA ) YEDA RES & DEV CO LTD.
PI   Ehrhardt G,  Kemper O,  Kuhnert P,  Wallach D;
DR   WPI; 96-010683/01.
DR   P-PSDB; W00332.
PT   Promoter regions of the human p75 tumour necrosis factor receptor
PT   gene  - and related transcription inhibitory region and sequence
PT   motifs, useful for inhibiting effects of TNF
PS   Claim 6; Figure 2B; 45pp; English.
CC   Sequence motifs and/or motif regions found within the promoter of
CC   p75 tumour necrosis factor receptor (TNF-R) can be used as probes to
CC   screen a human genomic or cDNA library to isolate factors that bind
CC   to them.  Inhibition/induction of such factors should allow
CC   modulation of TNF-R production and thus TNF binding.  In particular,
CC   they should allow inhibition of the deleterious effects of TNF which
CC   is involved in inflammation and is a mediator in septic shock, graft
CC   vs. host reactions, rheumatic and other autoimmune diseases.
SQ   Sequence   1827 BP;    391 A;    472 C;    542 G;    422 T;


  Query Match             17.8%;  Score 26.2;  DB 1;  Length 1827;
  Best Local Similarity   60.6%;  Pred. No. 12;
  Matches   43;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy     20 gactcagtgcttctggaggccaacaggacttctgagtcatcctgtgggggtggaggtggg 79
          || |||||||||||   ||  ||| ||||   |     |   ||||||  | ||| | |
Db   1553 GAGTCAGTGCTTCTGCCTGCAGACATGACTAGGGTACACTGAGGTGGGGAGGCAGGGGAG 1612

Qy     80 acaagggaaag 90
            ||  ||| |
Db   1613 GTAAAAGAAGG 1623


RESULT  10
V41229/c
ID   V41229 standard; DNA; 142 BP.
AC   V41229;
DT   01-OCT-1998  (first entry)
DE   Prevotella nigrescens target internal sequence (3' end) used as a probe.
KW   Bacteroides forsythus; Porphyromonas gingivalis; Prevotella intermedia;
KW   Prevotella nigrescens; periodontal disease; periodontopathic bacteria;
KW   AP-PCR; probe; ss.
OS   Prevotella nigrescens.
PN   US5789174-A.
PD   04-AUG-1998.
PF   11-OCT-1996; 729447.
PR   11-OCT-1996; US-729447.
PA   (UYLA-) UNIV LAVAL.
PI   Guillot E, Menard C, Mouton C;
DR   WPI; 98-446071/38.
PT   Probes, primers and methods for detecting bacteria - associated with
PT   periodontal disease
PS   Claim 19; Columns 17-18; 29pp; English.
CC   This represents a target internal sequence of Prevotella nigrescens
CC   that can be used as a probe for the specific detection of Prevotella
CC   nigrescens. The invention provides a composition of probes for the
CC   specific detection and identification of bacteria associated with
CC   periodontal disease, where the probes are obtained from amplification of
CC   nucleic acid sequences from bacteria selected from Bacteroides forsythus
CC   strain ATCC 43037, Porphyromonas gingivalis strain ATCC 33277, Prevotella
CC   intermedia strain ATCC 25611 and Prevotella nigrescens strain NCTC 9336,
CC   with one of the respective primers shown in V41210 to V41212. The probes
CC   are capable of selectively hybridising, under stringent hybridizing
CC   conditions, with the nucleic acid sequences from the specified bacteria.
CC   These probes are used in a method of detecting bacteria associated with
CC   periodontal disease in a sample obtained from the mouth of a human
CC   patient. Primer pairs (V41213 to V41220) derived from the nucleic acid
```

```
CC   sequences of the probes can also be used for the detection of the
CC   bacteria [sic] associated with periodontal disease. The probes can detect
CC   the presence of periodontopathic bacteria. The probes and primers provide
CC   a quick and specific identification without a need for the initial
CC   knowledge of the sequence of the targetted DNA.
SQ   Sequence   142 BP;    31 A;    34 C;    31 G;    44 T;


  Query Match             17.7%;  Score 26;  DB 1;  Length 142;
  Best Local Similarity   70.0%;  Pred. No. 7.2;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy     60 cctgtgggggtggaggtgggacaagggaaaggggtgaatggtactgctga 109
          ||||| ||| ||||   || |||||| || ||||||   || ||  | ||
Db    113 CCTGTCGGGTTGGAAAAGGCACAAGGTAAGGGGGTGGCAGGCACGACAGA 64


RESULT  11
V02016/c
ID   V02016 standard; cDNA to mRNA; 2525 BP.
AC   V02016;
DT   18-JUN-1998  (first entry)
DE   DNA encoding tumour antigen protein.
KW   Tumour antigen protein; gastric cancer; intracellular digestion;
KW   bind; major histocompatibility complex class I antigen; recognition;
KW   T-cell; gene therapy; tumour; autoimmune disease; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   5'UTR           1. .38
FT                   /*tag= a
FT   CDS             39. .2441
FT                   /*tag= b
FT   3'UTR           2439. .2506
FT                   /*tag= c
FT   polyA_signal    2507. .2527
FT                   /*tag= d
FT                   /note= "polyA site"
PN   WO9746676-A1.
PD   11-DEC-1997.
PF   04-JUN-1997; J01893.
PR   25-NOV-1996; JP-330424.
PR   07-JUN-1996; JP-168429.
PR   08-OCT-1996; JP-287572.
PA   (ITOH/) ITOH K.
PI   Imai Y, Itoh K, Shichijo S;
DR   WPI; 98-042184/04.
DR   P-PSDB; W44003.
PT   DNA encoding tumour antigen protein, fragments of which bind to MHC
PT   class I antigens - useful in gene therapy and auto:immune diseases
PS   Claim 2; Pages 40-42; 49pp; Japanese.
CC   The present sequence encodes a tumour antigen protein. It was isolated
CC   from a gastric cancer cell line. The tumour antigen protein has the
CC   ability to form fragments by intracellular digestion which bind to major
CC   histocompatibility complex (MHC) class I antigens to form a complex
CC   which is recognised by T-cells. The DNA is useful for gene therapy of
CC   tumours and autoimmune diseases.
SQ   Sequence   2525 BP;    625 A;    653 C;    906 G;    341 T;


  Query Match             17.6%;  Score 25.8;  DB 1;  Length 2525;
  Best Local Similarity   53.5%;  Pred. No. 18;
  Matches  54;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy     45 ggacttctgagtcatcctgtgggggtggaggtgggacaagggaaaggggtgaatggtact 104
          || ||||||||| | ||   |     |||||  ||| || ||      || |  ||   ||
Db   2381 GGTCTTCTGAGCCTTCTGCTTCTCCTGGAGCAGGGCCACGGTGCCCAGGGGCGTGTCGCT 2322

Qy    105 gctgattacaacctctggtgctgcctccccctcctgtttat 145
          |  | | |    ||   |    |||||| | ||| | || |
Db   2321 GGAGCTCATCTTCTTCAGGAGCGCCTCCTCGTCCAGCTTCT 2281


RESULT  12
Q61216
ID   Q61216 standard; DNA; 383 BP.
AC   Q61216;
DT   16-MAR-1994  (first entry)
DE   Human brain Expressed Sequence Tag EST01251.
KW   Gene transcription product; genetic markers; tagging; in vivo;
KW   transcription; mapping; locations; chromosomes; chromosomal; ss.
OS   Homo sapiens.
PN   WO9316178-A.
PD   19-AUG-1993.
PF   12-FEB-1993; U01294.
PR   12-FEB-1992; US-837195.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICE.
PI   Adams MD, Moreno RF, Venter CJ;
DR   WPI; 93-272882/34.
PT   Enriched oligonucleotides and corresp. sequences - used as
PT   markers for human genes transcribed in-vivo, facilitate tagging
PT   of most human genes
PS   Example 4; Page 456; 500pp; English.
CC   The Expressed Sequence Tag was isolated from a human brain cDNA
CC   library as part of a large set of ESTs which can be used as markers
CC   for human genes transcribed in vivo. They can be used to facilitate
CC   tagging of most human genes, for mapping locations of expressed genes
CC   on chromosomes, for individual or forensic identification, for mapping
CC   locations of disease-associated genes, for identification of tissue
CC   type, and for prepn. of antisense sequences, probes and constructs.
CC   EST01251 has a "good" coding probability as evaluated using the
CC   coding-region prediction program CRM. See also Q59041-Q61440.
SQ   Sequence   383 BP;    118 A;    90 C;    94 G;    80 T;


  Query Match             17.4%;  Score 25.6;  DB 1;  Length 383;
  Best Local Similarity   54.2%;  Pred. No. 13;
  Matches  52;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy     12 ccaaccatgactcagtgcttctggaggccaacaggacttctgagtcatcctgtgggggtg 71
          || ||| |    | ||   |    || |||  |    |   |   || ||  | ||| | |
Db    169 CCTACCTTTGGGCGGTCTATACTGAAGCCCTCCACTCCCATCCCTCCTCAAGAGGGTGAG 228

Qy     72 gaggtgggacaagggaaaggggtgaatggtactgct 107
          |||||||||  |  | |  | || | | | || ||||
Db    229 GAGGTGGGAGAGAGCAGTGAGGAGCAAGATAATGCT 264


RESULT  13
Q58718
ID   Q58718 standard; DNA; 1254 BP.
AC   Q58718;
DT   30-SEP-1994  (first entry)
DE   Human gamma-1 chain first membrane exon.
KW   Membrane; exon; human; gamma-1; chain; membrane anchoring peptides;
KW   heavy chain; isotype; cell surface; extracellular region;
KW   migis segment; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   exon            547. .678
FT                   /*tag= a
PN   US5298420-A.
PD   29-MAR-1994.
PF   03-AUG-1990; 562201.
PR   03-AUG-1990; US-562201.
PR   19-JUN-1992; US-902449.
PA   (TANO-) TANOX BIOSYSTEMS INC.
PI   Chang TW;
DR   WPI; 94-100338/12.
DR   P-PSDB; R60128.
PT   Monoclonal antibodies and fragments specific for extracellular
PT   migis epitope(s) on B cells - used to bind B cells for
PT   immunosuppressive purposes or to remove B cells from the
PT   circulation
PS   Disclosure; Fig 1A; 13pp; English.
CC   The sequences given in Q58718-19 represent the first and second
CC   membrane exons of the human gamma-1 chain which encode membrane
```

```
CC   anchoring peptides.  Membrane anchoring peptides are fragments of
CC   the heavy chains of the associated isotypes and anchor the isotypes
CC   to the cell surface.  The extracellular region of the encoded gamma-1
CC   chain contains numerous acidic residues and is thought to comprise
CC   the migis segment.
SQ   Sequence   1254 BP;    218 A;    515 C;    281 G;    240 T;


  Query Match             17.4%;  Score 25.6;  DB 1;  Length 1254;
  Best Local Similarity   66.1%;  Pred. No. 17;
  Matches   37;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      59 tcctgtgggggtggaggtgggacaagggaaaggggtgaatggtactgctgattaca 114
           |||| |||| |   || ||||||  |||| | |||     ||| |||||||   ||
Db     288 TCCTCTGGGCCTCAGGGAGGGACACGGGAGACGGGGAGCGGGTCCTGCTGAGGCCA 343


RESULT  14
Q47076/c
ID   Q47076 standard; DNA; 1371 BP.
AC   Q47076;
DT   13-JAN-1994  (first entry)
DE   B. canis 21B4/rhoptry antigen gene 2 DNA.
KW   Polymerase chain reaction; PCR; amplify; primer; detection;
KW   babesiosis; parasite; Babesia bovis; 21B4/rhoptry; antigen; gene;
KW   repeat region; immune response; vaccine; ss.
OS   Babesia canis.
PN   WO9314204-A.
PD   22-JUL-1993.
PF   15-JAN-1993; AU0012.
PR   15-JAN-1992; AU-000399.
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PI   Dalrymple BP,  Peters JM;
DR   WPI; 93-243219/30.
DR   P-PSDB; R39902.
PT   Detecting closely linked gene copies which encode protective
PT   antigen against babesiosis - by screening babesial genomic DNA
PT   library with oligo-nucleotide probe based partial sequencing of
PT   protective antigen and identifying positive clones
PS   Claim 24; Fig 7; 55pp; English.
CC   This sequence represents the Babesia canis 21B4/rhoptry antigen gene
CC   2.  This sequence was determined from restriction fragments from the
CC   clone B. canis lambda GEM-11 #9.  B. canis was found to contain two
CC   genes which are related to the B. bovis 21B4 gene.  Gene 1 and gene 2
CC   are very similar but gene 2 appears to contain a large number of
CC   repeats.  Babesia antigen genes can be used in the production of a
CC   combined vaccine which will stimulate a greater immune response and
CC   afford broader immunity than a single antigen vaccine.  See also
CC   Q47068-74.
SQ   Sequence   1371 BP;    418 A;    312 C;    325 G;    318 T;


  Query Match             17.3%;  Score 25.4;  DB 1;  Length 1371;
  Best Local Similarity   58.7%;  Pred. No. 21;
  Matches   44;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy      45 ggacttctgagtcatcctgtgggggtggaggtgggacaagggaaaggggtgaatggtact 104
           ||  |||  | ||  |||   |     | || |||| |||| ||    || |||| | ||
Db     696 GGCGTTCCTAATCTGCCTTCCGAACAAGTGGCGGGAGAAGGCAATTTTGTTAATGATCCT 637

Qy     105 gctgattacaacctc 119
           |    ||||||| ||
Db     636 GGCATTTACAACATC 622


RESULT  15
N70930
ID   N70930 standard; cDNA; 565 BP.
AC   N70930;
DT   21-MAY-1991  (first entry)
DE   Sequence encoding protein with the biological activity of HUSI
DE   (human seminal plasma inhibitor) type I inhibitors encoded on pRH 34.
KW   Bronchitis therapy; cervix inflammation; hyperfibrinolysis; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             59. .457
FT                   /*tag= a
PN   DE3600571-A.
PD   06-AUG-1987.
PF   10-JAN-1986; 600571.
PR   10-JAN-1986; DE-600571.
PA   (CHEF ) GRUNENTHAL GMBH.
PI   Heinzel R, Appelhans H, Gassen HG, Machleidt W, Seemuller U;
DR   WPI; 87-222038/32.
DR   P-PSDB; P70584.
PT   New protein with human seminal plasma inhibitor activity - and
PT   new DNA coding sequences, vectors and transformed organisms,
PT   useful e.g. for treating bronchitis or inflammation
PS   Claim 5; Fig 5; 28pp; German.
CC   HUSI-type I inhibitors are useful for treating chronic bronchitis,
CC   chronic cervical inflammation; inflammation associated with
CC   excessive mucus prodn.; post-operative bleeding caused by
CC   hyperfibrinolysis, and shock. They are esp. suitable for
CC   administration as spray or inhalation. The expression control
CC   sequence on the DNA fragment in N70928 is bound to the 5' end of
CC   HUSI cDNA in an expression vector.
SQ   Sequence   565 BP;    106 A;    162 C;    171 G;    126 T;


  Query Match             17.1%;  Score 25.2;  DB 1;  Length 565;
  Best Local Similarity   57.7%;  Pred. No. 19;
  Matches   45;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy      65 gggggtggaggtgggacaagggaaaggggtgaatggtactgctgattacaacctctggtg 124
           ||||| || || |||    |||   |||| |   |  ||| |||  | || | |   ||
Db       7 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGTCACTCCTGCCTTCACCATGAAGTC 66

Qy     125 ctgcctccccctcctgtt 142
           | ||  || | |||  ||
Db      67 CAGCGGCCTCTTCCCCTT 84


Search completed: June 24, 1999, 01:22:20
Job time: 2656 sec
```